

AUTO3DEM Modification History

auto3dem_v3.02

- Overhauled parallelization scheme for programs PPFT, PO2R, and P3DR. Reduced memory requirements in PPFT by approximately 40%. Test problems run on 16 processors showing reductions in run time of 25-40% relative to v3.01.
- Utility program autopp added for performing repetitive tasks such as converting files from one format to another or globally replacing a string in a set of files. Note that not all are currently operational. This script should still be considered as under development.
- Fixed bug in PO2R that resulted in ctfmode being hard coded to 1.
- Defocus values in particle parameter files only tested if CTF corrections are used.
- Header always written to auto3dem summary file if file does not already exist, regardless of value of restart flag.
- Minor bug fix to PO2R so that parameter nangle specifying number of steps to be taken along each orientation angle can be set to zero. Doing this allows PO2R to refine particle image origins while leaving orientations unchanged.
- Modify tiff2pif conversion program to optionally accept third command line argument that specifies binning factor. Formerly, users were prompted to enter value.
- Add mrc2pif, pif2mrc, and pif2ccp4 conversion programs. Modify so that programs can either accept file names on the command line or prompt user to enter names.
- Provide options for building applications using static linking.
- Executables now reside in bin/ directory rather than BIN/. To ensure back compatibility, BIN is now a symbolic link that points to bin/.
- Improved instructions provided on using “ppft verbose -1” option in auto3dem input file that is automatically generated by setup_rmc.
- Move perl modules (.pm files) into modules/ directory and modify perl scripts to look for modules in new location. Replace “use lib do” construct with “use FindBin”.
- Add documentation files that are used by programs robem, oned, ctfDisp, and fixpif.
- Minor cleanup of argument list for subroutine global_cc in program PPFT. Also affects calling routine global.

auto3dem_v3.01

- Minor modifications to make files to ensure that applications build properly when using Portland Group compilers. The -Mnomain option is now used only with those applications that require access to Fortran libraries, but have main program declared in C source. Remove getarg subroutine from libCommpk to avoid clashes with Portland Group library function. Declare iargc to be of type integer in routine parse_cmd_line.
- Add mrc2pif and pif2mrc conversion routines to convert directory.
- Modify calls to XtVaSetValues() from function defocusArrowActivate() in ctfDisp program to avoid segmentation faults on 64-bit machines.
- Cleanup preprocessor directives in robem source.

auto3dem_v3.0

This release gathers for the first time all of the image reconstruction, image preprocessing, and utility codes into a single archive. New additions include robem, ctfDisp, oned, fixpif, tiff2pif, emmap3dt, emmapzoom, emsf, em3dbt, emprj, diffit, normit, zerodens, and a number of conversion routines. The make_all script now takes two command line arguments

```
make_all (parallel | serial) (all | gui | nogui)
```

The first argument behaves exactly as before, while the second argument specifies which set of applications should be built. In most cases, the second argument should be 'all', but the 'nogui' option may be chosen when doing a build on a machine that does not have the Motif library installed. The 'gui' option is generally not needed, but is provided for added flexibility.

If you are already a user of robem, please be aware that this release contains a major overhaul of the software. Modifications include:

- All Fortran code has been brought up to the Fortran90 standard
- Much redundant, obsolete, or unused code has been deleted
- Old Kernighan and Richie (K&R) style function prototypes have been replaced with prototypes that conform to the ANSI C standard
- Directory structure, make files, and include files have been cleaned up and reorganized
- Minor bugs, mostly related to type mismatches in C function calls, have been implemented
- Routines used by multiple applications have been moved into libraries

In order to build many of the newly added programs, you will need to install the Motif library. We do not currently have a feature in place for automatically determining whether the code is being compiled on a 32-bit or 64-bit architecture. The make.inc.common file is hardcoded for 64-bit hardware and the MOTIFLIB macro will need to be manually edited if you are running on a 32-bit machine.

Running the new programs requires that the EMDIR environment variable be set. The easiest way to do this is to set the EMDIR variable in the .cshrc file and then use it to append the path variable.

```
setenv EMDIR /path_to_programs/BIN
set path = ($EMDIR $path)
```

All make files have been overhauled and a number of subroutines that are common to both the program PPFT (parallel PFTsearch) have been moved into libCommpk. In addition, several of the routines in the PPFT directory have been renamed so as to avoid conflicts with other routines of the same name in libEMF.

| | | |
|--------------|---|-------------------|
| fft_map_fill | → | fft_map_fill_ppft |
| map_fft_fill | → | map_fft_fill_ppft |
| map_prj | → | map_prj_ppft |
| map_sym_cavg | → | map_sym_cavg_ppft |

The following directories have been created to accommodate the newly added code

convert
conway_tiff2pif
ctfDisp
em_tools

fixpif
lib3DAll
libEMF
oned

pftprj
robem

auto3dem_v2.04

- Overhaul make files and define library macros in include files. Rename libraries to use more standard naming conventions and rename directories as follows to be more consistent with other Purdue software

| | |
|--------|--------------|
| PIFlib | → libR |
| Commpk | → libCommpk |
| Compar | → libCompar |
| DIElib | → libDIERCKX |
| Vfftpk | → libVfftpk |

- Add new script `flesh_out.pl` to generate complete particle parameter files from files that contain only header lines.
- Fix minor bug in `select.pm` related to filtering particles on theta and phi
- Get rid of old code delimited by `#define OLDWAY` preprocessor directives in `libPIF.c`
- Generate new `libPIF.h` header file from `libPIF.c`

auto3dem_v2.03

Important notes for building auto3dem

Auto3dem and the parallel codes that it calls can now be built and run in either serial or parallel mode. An implementation of the MPI library (*e.g.* mpich) is no longer needed when running auto3dem on a single processor. The `make_all` script used to build the executables now takes a single command line argument, with allowed values of 'parallel' or 'serial'. Symbolic links are set by the script as follows

```
% make_all parallel
% ls -g make.inc BIN/mode.pm
lrwxrwxrwx 1 csd357 14 Jun 13 15:14 BIN/mode.pm -> mode_parallel.pm
lrwxrwxrwx 1 csd357 15 Jun 13 15:14 make.inc -> make.inc.parallel

% make_all serial
% ls -g make.inc BIN/mode.pm
lrwxrwxrwx 1 csd357 14 Jun 13 15:14 BIN/mode.pm -> mode_serial.pm
lrwxrwxrwx 1 csd357 15 Jun 13 15:14 make.inc -> make.inc.serial
```

When building auto3dem for parallel operation, the serial Fortran 90 and ANSI C compilers are called by the corresponding `mpif90` and `mpicc` scripts. For builds in serial mode, the FC and CC macros in the `make.inc.serial` file must be manually edited if you are not using the `gfortran` and `gcc` compilers. The `make_all` script determines whether or not a build of auto3dem already exists. If the mode (serial or parallel) of the current build differs from that of the previous build, all .F files are 'touched' to ensure that they are run through the preprocessor and recompiled.

For auto3dem runs in serial mode, the number of CPUs no longer needs to be specified and any value set using the `-ncpu` flag is ignored.

Detailed listing of code changes:

- auto3dem.pl and run_mpi_prog.pm modified to use mode.pm module. When running in serial mode, auto3dem.pl no longer requires (and quietly ignores) specification of number of nodes. validate.pm module modified to accept mode argument. Version string now specifies whether serial or parallel build is being used.
- make_all script modified to accept mode (serial or parallel) on command line and set symbolic links to appropriate make include files and Perl modules before initiating build. Make include files simplified to just make.inc.parallel and make.inc.serial.
- Argument intents added to P3DR/exchange_2_slab.f
- File extension for P3DR/realtocomplx.F and matrixentries_slab.F changed to .f
- Makefiles for POR, PCTFR, PSF, PPFT, and PCUT modified to include generic rule for creating objects from Fortran source with .F extension.
- Add preprocessor directives to following files to isolate parallel code and allow optional builds in either serial or parallel mode. Add serial code where needed to replace functionality carried out by MPI routines (*e.g.* initialization of process number and number of nodes, data reduction). All files with .f extension renamed to use .F extension so that they will be recognized by preprocessor.
 - Compar/bcast_parameters.f
 - Compar/error_stop.f
 - Compar/exch_3d_1.f
 - Compar/gather3d.f
 - Compar/output_density.f
 - Compar/read_files.f
 - Compar/read_map.f
 - P3DR/P3dr.F
 - P3DR/cmpt_intrps.F
 - P3DR/exch_intp.F
 - P3DR/exchange_2_slab.f
 - PCTFR/Pctfr.f
 - PCTFR/cmpt_ctf.f
 - PCUT/Pcut.f
 - PCUT/cut_map.f
 - POR/Por.f
 - POR/cmpt_ort.f
 - PPFT/global.f
 - PPFT/global_cc.f
 - PPFT/pftsearch.f
 - PSF/Psf.f
 - PSF/comp_sfactor.f
- Cleanup and modernization of routines in PPFT

- Makefile – remove map_peak_rest.o from object list
- map_peak_restr.f – remove; functionality now in pftcc_peak
- list_ccs.f - cleanup and intent specification
- pftcc_peak.f - major cleanup and reorganization for clarity; move call to rearrange_fft_style from here to get_xy; replace call to map_peak_rest with inlined functionality.
- get_xy.f - add call to rearrange_fft_style after call to ccf_fft
- calc_pfts_g - add FILT_FAC to argument list; declare pfft to be complex rather than real and make necessary code changes
- tsbend.f - general cleanup and intent specification
- rearrange_fft_style.f - rewritten so that rearrangement of array done in place. Single array now passed with intent INOUT and same dimensions as array in calling routine.
- pftsearch.f - get rid of variables NPHI and NTHE
- global.f - get rid of variables NPHI and NTHE; specify argument intents
- key_info.f - get rid of variables NPHI and NTHE, specify argument intents and completely overhaul routine to improve clarity and readability
- calc_mod_tps.f - rename (NPH, NTH) to (NPHI, NTHE) and declare as local variables rather than subroutine arguments.
- get_nview_mod.f - declare NTHE as local variable rather than subroutine argument.
- pmap_fft.f - declare array A to be complex of size NROT rather than real with size 2*NROT and make corresponding code modifications.
- fft_2d.f - specify argument intents and cleanup comments.
- fft_2d_back.f - specify argument intents, cleanup comments, and use Fortran 90 array syntax.
- global_cc.f - declare array PPFT to be complex rather than real; delete flip, index, mode, and sql from argument list in call to get_tpo_g; move hand flip calculations from here into get_tpo_g.
- get_tpo_g.f - specify intent for all subroutine arguments; delete flip, index, mode, and sql from argument list; move hand_flip calculations here from global_cc; declare pfft to be complex and make necessary code modifications
- ccf_fft - Complete overhaul for improved clarity, use of complex arrays, cleanup of logic, commenting. In particular, switching from real to complex arrays vastly simplifies index calculations.
- fill_params.f - specify argument intents, extensive overhaul of logic for improved clarity
- write_params.f - overhaul and cleanup for improved clarity; specify argument intents.

auto3dem_v2.02

Summary of changes that result in slight numerical differences

- Faster approximate algorithm for determining ω in program PPFT. Enabled by default and controlled by 'ppft quick_omega' parameter.
- Faster approximate algorithm in POR for local search of orientation space. Enabled by default and controlled by 'por quick_search' parameter.
- Apodized map produced by P3DR. Width of border region controlled by 'p3dr apo_border' variable, default value equal to 12 pixels.
- Automatic refinement of CTF parameters now turned off by default. Controlled by 'auto refine_ctf' parameter.

- Bug fix in auto3dem so that inner and outer radii of capsid estimated correctly when binning used in PPFT

Detailed description of changes

- Major performance enhancements made to PPFT routine (get_phiomega) responsible for determining the orientation angle ω and the sign of ϕ . For smaller problems, this routine accounts for a very tiny fraction of total run time, but owing to the scaling behavior of the algorithm (NROT³, where NROT depends on map size and binning factor), get_phiomega can dominate the run time for larger problems.

The initial search for ω is now done over a coarser grid of values, with the step size dependent on NROT. This is followed by a local search using a finer grid in the vicinity of the top scoring values of ω obtained during the initial search. After several iterations of global search mode, only a small fraction of the particle images have orientations that differ from those obtained using the original algorithm and the actual differences in the orientations are minimal.

The “quick omega” feature is enabled by default in auto3dem. Adding the following line to auto3dem input files disables this option

```
ppft quick_omega 0 # auto3dem input file
```

This new capability required the creation of a new subroutine get_phiomega_quick and modifications to the following files in directory PPFT: global.f, global_cc.f, get_tpo_g.f, pftsearch.f, ppft_info.f, and key_info.f. Also required changes to include/infohead.inc, init_params.pm, and make_program_input.pm

- Capabilities added to P3DR to produce an apodized map, with a Gaussian falloff applied to the density in the region $\text{boxrad-border} \leq r \leq \text{boxrad}$. Border has units of pixels and must be an integer. Using a border width of zero recovers original P3DR behavior. Default value in auto3dem set to 12, but can be overridden by adding the following line to the auto3dem input files

```
p3dr apo_border n # auto3dem input file
```

This new capability required changes to the following files: P3DR/P3dr.F, P3DR/density_clear.f, Commpk/info.f, Compar/bcast_parameters.f, info.inc, init_params.pm, and make_program_input.pm

- Provide “quick search” capabilities in POR so that a restricted local search of orientation space (one $\Delta\theta$ step along each direction for each orientation angle) is performed first and only those particle images that find a better orientation in this restricted region are subjected to a more extensive local search. Using the “quick search” feature typically results in a very small fraction of the particle images settling into orientations that are different from those obtained using the more extensive local search, but has been shown to decrease run times by up to a factor of five.

The “quick search” feature is enabled by default in auto3dem. Adding the following line to auto3dem input files disables this option

```
po2r quick_search 0 # auto3dem input file
```

This new capability required changes to the following files: POR/por.f, POR/cmpt_ort.f, Commpk/info.f, Compar/bcast_parameters.f, info.inc, init_params.pm, and make_program_input.pm

- Refinement of CTF parameters using program PCTFR is now disabled by default. To override, add the following line to auto3dem input file

```
auto refine_ctf 1
```

- setup_rmc.pl prints comments to the auto3dem input file describing how to set parameters to perform faster reconstructions and reach higher resolutions. The following commented out parameters are now listed
 - ppft bin_factor
 - ppft verbose
 - ppft annulus_low
 - ppft annulus_high
 - pcut in_rad
 - pcut out_rad
 - auto freeze_annulus
 - auto bin_reduce
- Bug fix made to Perl module get_ann_lo_hi.pm so that inner and outer radii of capsid are estimated correctly when binning is used in program PPFT.
- POR modified to keep statistics on both the level of individual micrographs and the entire run for the number of particle images that fall into each of the following categories
 - Stable orientation / stable origin
 - New orientation / stable origin
 - Flipped hand / stable origin
 - Stable orientation / new origin
 - New orientation / new origin
 - Flipped hand/ new origin

The “stable” designation means that the orientation or origin for a particle is unchanged after running POR. “Flipped hand” means that the orientation is the same except for a change in the handedness of the particle. This information is summarized in the POR output file.

Extra fields have been added to the new particle parameter files generated by POR so that the status of the orientation (Stable, New, Flipped) and origin (Stable, New) is listed for each particle.

- Add new Perl module to auto3dem to perform tests that determine whether or not particle parameter files are properly formed. Affects files auto3dem.pl and sanity_checks.pm
- Allow for specification of P3DR parameters magfactor (default value 1.0) and map_dim (default value 0) in auto3dem input file. Affects files init_params.pm and

make_program_input.pm

- Modify PCTFR/cmpt_ctf.f to handle arbitrary formats in the particle records of the particle orientation files. Records are now read in as strings, rather than parsed into individual fields, and written out to the new particle file following the updated CTF parameters. This allows us to include additional fields that may be written by programs POR or PPFT without having to make modifications elsewhere in the code.
- Command line syntax for auto3dem and setup_rmc more flexible. Key value pairs no longer need to be separated by equal signs. For example, both of the following are now valid

```
%auto3dem -ncpu=16 -input=input_file  
%auto3dem -ncpu 16 -input input_file
```

- Minor performance enhancements made to $(\theta, |\phi|)$ search routine get_thephi in program PPFT.
- Make following modifications to PPFT input routine ppft_info:
 - Equal sign in key-value pairs is now optional
 - Strings containing '/' properly handled
 - lg_out initialized to zero
 - Start of data file list can be specified using both 'end_of_keys' (first three characters significant) or 'inputparameterfiles' (first nine characters significant).
- Allow blank lines and comments (using '#') in the data file section of input for programs PPFT, POR, P3DR, and PCTFR. Required changes to PPFT/pfile_info_pft.f and Commpk/pfile_info.f.

The remaining modifications listed below for v2.02 all involve internal code changes that do not affect results, use of codes, or specification of input parameters.

- Cleanup timing functionality throughout code
 - Add new function elapsed_time.f to library Commpk that calculates the time elapsed between pairs of system_clock readings; get rid of old timer function timer1.c, and update Makefile.
 - Modify mem_time.inc so that all timing variables reside in a common block and get rid of test_memory, test_time, and wtime declarations.
 - Replace calls to mpi_wtime with Fortran 90 system_clock intrinsic in the following files: Ctffit/ctffit.f, P3DR/P3dr.F, P3DR/cmpt_intrps.F, PCTFR/Pctfr.f, PCTFR/cmpt_ctf.f, PCUT/Pcut.f, POR/Por.f, POR/cmpt_ort.f, PPFT/global.f, PPFT/global_cc.f, PSF/Psf.f
- Get rid of old commented out code in following files: P3DR/matrixentries_slab.F, PCTFR/Pctfr.f, POR/Por.f, POR/cmpt_ort.f.
- Comment out or remove unused variables and labels in following files: PCTFR/cmpt_ctf.f, PPFT/list_ccs.f, PPFT/pfile_info_pft.f, PPFT/ppft_info.f, Ctffit/ctffit.f, Ctffit/rotate.f

auto3dem_v2.01

- Modify programs PSF, PCUT, P3DR, PCTFR, PPFT and POR to allow specification of input file both as command line argument and through redirection of standard input. For example, both of the following are valid ways to execute P3DR

```
% mpirun -np 8 P3DR p3dr_input > p3dr_output
% mpirun -np 8 P3DR < p3dr_input > p3dr_output
```

The first syntax is preferred and is used internally by auto3dem when launching the parallel jobs. The reason for making this change is that many MPI implementations cannot handle redirection from standard input if the number of characters is greater than 4096. **The second syntax is retained for continuity purposes only and is no longer recommended.**

- Add full path to nodefile in setup_rmc and auto3dem.
- Add capabilities to change search mode bin factor in auto3dem using tests analogous to those used in making decision to switch from search mode to refine mode. Controlled through 'auto bin_reduce' parameter. Bin factor is now listed in summary file.
- Fix logic error in auto3dem that resulted in mode (search → refine) being updated after the restart file was written. This bug did not affect results of image reconstruction, but could result in an extra iteration of search mode being carried out if the calculations died during the first iteration following the automated transition from search to refine mode.
- Added new script config_test.pl to main code directory to perform simple tests on the computing environment. Display Perl version is displayed and determine whether or not all required and optional Perl modules are present. Confirm that mpirun, mpicc, mpif90, and scalar C and Fortran compilers called by mpif90 and mpicc, respectively, are in path.

The remaining modifications listed below for v2.01 all involve internal code changes that do not affect results, use of codes, or specification of input parameters.

- Eliminated temporary array PRJSL in file PPFT/global.f
- Cleanup and reorganization of PIFlib/libPIF.c

auto3dem_v2.0

Important - Version 2.0 contains a new command line interface for running setup_rmc and auto3dem. The old usage syntax is no longer valid. Any scripts that call setup_rmc or auto3dem must be modified to use the new syntax.

- Completely overhauled command line interface for both auto3dem and setup_rmc so that input is done using key-value pairs. For example

```
setup_rmc -ncpu=4 -seed=123 -list=listfile
setup_rmc -usedefaults
```

```
auto3dem -ncpu=12 -input=input_file -nodefile=mynodelist
```

setup_rmc can be run purely with default values using the -usedefaults flag. For

auto3dem, values must be provided for the number of CPUs and the name of the input file. The nodefile no longer needs to be specified for either setup_rmc or auto3dem. This is to be contrasted with previous versions where the word “none” had to be explicitly used for those cases where the parallel computing environment did not require a node list. For batch systems running the PBS scheduler, the PBS node file is automatically obtained from the \$PBS_NODEFILE environment variable via Perl’s %ENV hash.

Both programs list usage information if executed either without any arguments or with the -help flag. All flags are case insensitive and whitespace before or after the “=” in the key value pairs is tolerated. In addition, setup_rmc prompts the user to continue run using all default values after printing usage information.

- Major performance improvements in P3DR for icosahedral symmetry. Twofold and threefold symmetry operations are applied to 3D DFT of model, thereby reducing the number of interpolations required for each image from 60 to 5. Speedup is problem and system dependent since it depends on the number of particle images and the relative time required for different operations (FFT, interpolation, etc.), but runs times have been measured to be reduced by 8x for test cases involving 1000 particle images.
- Added capability to handle both relative and absolute path names to boxed image files (as specified in particle parameter files) to setup_rmc.
- CTF corrections can be turned off globally using -noctf option on the setup_rmc command line and “auto noctf 0” in the auto3dem input file. This has the same effect as manually setting the CTF mode for programs P3DR, PO2R, PCTFR, and PPFT to zero. Using the -noctf option in setup_rmc automatically propagates setting of the noctf flag in the generated auto3dem files.
- The auto3dem input file now recognizes the keys ctf_mode and ctfmode for all programs that perform CTF corrections. This change was made to remedy the confusing situation where PPFT used ctf_mode while P3DR, PO2R, and PCTFR used ctfmode.
- Intermediate files generated during random model method calculations are all moved to a new directory named RMC_temp. The best starting model is renamed rmc.pif and moved into the directory containing the particle parameter files. Files and directories corresponding to random orientations that are not used are automatically deleted using a new utility script, remove_useless.pl, which is called from RMC_run. A short script named RMC_cleanup is generated by setup_rmc for the purpose of recursively removing contents of RMC_temp.
- setup_rmc automatically generates a basic auto3dem parameter file to continue reconstruction after random model method calculations have been completed.
- Sensible defaults now used for obtaining list of particle parameters files considered by setup_rmc. If -list key is not used, setup_rmc first looks for a file named ‘list’. If this file does not exist or is not readable, then the data directory is queried for files of the form *000, *001, etc.

When a list of parameter files is read from a file, now have added flexibility in the specification of the files. Wildcards, comments, blank lines, and whitespace are now

allowed. Can also use a single quoted expression in place of the file. For example, `-list='*001'` will be expanded within `setup_rmc` to a list of all files ending with `001` in the specified directory.

- Modified `auto3dem.pl` so that `GD::Graph` module is loaded at run time using “require” rather than “use”. This makes it possible to trap exceptions and bypass the FSC graph generation if the module cannot be found. The advantage of this approach is that the `auto3dem.pl` source does not have to be manually edited to handle Perl installations that are missing `GD::Graph`.
- PCTFR is now called by `auto3dem` only if ‘auto mode’ equals refine and both ‘pctfr ctfmode’ and ‘auto refine_ctf’ are true (non-zero).
- Lower memory algorithm and single precision interpolation option for P3DR have been removed. This both simplifies the source code and makes it easier to implement planned performance enhancements. The default is to use the more memory intensive (and less communications intensive) algorithm and double precision interpolation options, but these could previously be overridden using the `-DALG3` and `-DSPREC` compiler flags.

An older version of P3DR has been retained in the `P3DR_old` directory, but does not contain the performance enhancements described above.

- Semicolons now allowed, in addition to commas and whitespace, as delimiters in list of email addresses in `auto3dem` input file. For example,

```
auto recipient address1 address2, address3; address4
```

- MPI communications in `PPFT/global.f` modified to avoid exceeding MPI buffer limitations. Allows PPFT to be run using a finer spacing between projections of the model and/or larger models.
- MPI communications in `P3DR/exch_intp.F` modified to avoid exceeding MPI buffer limitations. Allows P3DR to be run for larger image sizes.
- `PSF/Psf.f` modified to give warning message rather than terminating execution if the pixel sizes in the even and odd map headers are different.
- `PCTFR/Pctfr.f` modified to use pixel size as specified in the particle parameter files rather than the map files. Warning messages are provided if the pixel size of the map is either negative or larger than the particle parameter file pixel size.

The remaining modifications listed below for v2.0 all involve internal code changes that do not affect results, use of codes, or specification of input parameters.

- Variable `r4_max` set to `huge(r4_max)` in `PCTFR/cmpt_ctf.f` and `POR/cmpt_ort.f`
- General cleanup and improved commenting of `Compar/bcast_parameters.f` and `Compar/read_files.f`

- Remove nodefile from validate_commandline_args() argument list in validate.pm and from call to function in auto3dem.pl. Allow nodefile to be undefined in auto3dem, writer_header, and run_mpi_prog.

auto3dem_v1.12

- Default values are now set for the “auto outfile” and “auto rundir” parameters. Unless otherwise specified, outfile is set to the name of the working directory where auto3dem is launched and rundir is set to be the directory dat. This involved changes to the file init_params.pm.
- Modified Fortran input routines pfile_info and pfile_info_pft to handle more flexibility in the specification of the image file names in the particle parameter files. The image name is first tested exactly as entered in the parameter file. If the program is unable to open it, the path name to the file is stripped off and an attempt is made to open the file in the current working directory. Similar changes made in the Perl module get_info.pm.
- Rename the auto3dem parameters fsc_locut and fsc_hicut to fsc_lothresh and fsc_hithresh, respectively. Scripts will still accept the old parameter names, but will be converted internally to new names. These changes were made so that the names of the variables would more accurately reflect their usage as thresholds in the FSC data. Similarly named variables in the Perl code were renamed to use the _lothresh and _hithresh suffixes. Also reordered return argument list from PSF parsing routine. These changes involved modifications to the files init_params.pm, auto3dem.pl, psf_parse.pm, and updatre_res.pm.
- Minor bug fixes to the files Commpk/ctf_para.f, POR/Por.f, and P3DR/P3dr.F to properly handle the case where no CTF correction is applied. PCTFR/Pctfr.f modified so that it exits early if ctfmode is set so that no correction is applied. Commpk/info.f fixed to properly handle num_den_pix set equal to zero in the input files.

The remaining modifications listed below for v1.12 all involve internal code changes that do not affect results, use of codes, or specification of input parameters.

- Intent of interpolants argument in P3DR/cmpt_interps.F changed to INOUT.
- In the following routines, replace the variable stdoutoutput with * in write statements: Commpk/intlz_arrays.f, Commpk/symmcode.f, P3DR/P3dr.F, P3DR/exchange_2_slab.f, P3DR/fftsynth_1_m_slab.f, P3DR/fftsynth_1a.f, PCTFR/Pctfr.f, PCUT/Pcut.f, and POR/Por.f
- Completely overhaul the include file include/allprog.inc. Get rid of stdinput, stdoutoutput, max_input, pi, twopi, deg_to_rad, and rad_to_deg; use Fortran90 parameter statements to set constants; improve comments; move filename_len from allprog.inc to info.inc. Get rid of variable skipone since it is used only in readorient.f and declared with SAVE attribute to retain value between calls.
- Cleanup and overhaul include file include/info.inc. The variable filename_len is now declared here so that the use of info.inc does not rely on allprogs.inc.

- Get rid of skipone variable in Commpk/info.f
- Replace filename_len with hardcoded value 257 in files Compar/read_map.f and Ctffit/ctffit.f
- Get rid of oddeven parameter in POR/cmpt_ort.f and PCTFR/cmpt_ctf.f; in calls to readorient, replace with zero.
- In routine Commpk/intlz_params.f get rid of initialization of pi, twopi, rad_to_deg, and deg_to_rad
- Overhaul logic in PPFT routine involved with the VAROPT feature. This is an experimental feature that is currently not recommended, but code modifications may have fixed a long time bug. Changes involved files global_cc.f, pftcc_fill_g.f, get_tpo_g.f, avg_pftimg.f, and get_thephi.f. Also declare variable pi locally in pftsearch.f

auto3dem_v1.11

- Email notification feature modified so that messages can be sent to multiple email addresses and that an arbitrary number of files can be attached. Individual addresses may be separated by whitespace and/or commas and each individual address is tested to be well formed. Note that additional fields in auto3dem parameters file records are still ignored except for the “auto recipient” record. Modifications involved files auto3dem.pl, sendmail.pm, and init_params.pm
- Capabilities added to generate and email graphs of FSC curves using GD library. Since not all sites have the required GD and GD::Graph Perl modules, the code in auto3dem.pl marked with the #GDSFC to construct the graphs must be manually uncommented.
- Input files info.f and pfile_info.f in library Commpk completely overhauled. Can now handle P3DR and PO2R input files that contain blank lines, tabs and leading whitespace.
- OPEN statements in key_info.f and write_params.f modified so that PPFT no longer dies if it tries to overwrite existing files.
- Added subroutine indexx (borrowed from Numerical Recipes in Fortran90, with very minor modifications) to the Commpk library. Will be used in future releases to sort records in particle parameter files by their ID.
- PCTFR/cmpt_ctf.f and POR/cmpt_ort.f remove all whitespace from particle image file names before writing particle parameter files.
- POR/cmpt_ort.f now writes zeroes for the last two fields in the particle parameter files. By making this change, the files generated by PPFT and POR will have the same format.
- Commpk/readorient.f modified so that it can deal more reliably with particle parameter files that are missing the last two scores. These types of files are generated by OOR and older versions of PO2R.
- File Commpk/read_1_pif.f cleaned up and commented. Include statements removed and argument list expanded. Corresponding changes made to argument list in Por.f, P3dr.f,

and Pctfr.f.

- Minor changes made to auto3dem so that:
 - Temporary files stdin_temp and message.txt are deleted
 - Log and summary files are properly handled in the event that a new run is launched but the log and/or summary files already exist
 - FSC data and graphs labeled using same prefix as map, summary, log, and restart files.

auto3dem_v1.10

- setup_rmc no longer specifies the number of models on the command line. New syntax is

```
setup_rmc dir list ncpu nodefile [boxrad]
```

- Streamlined random model calculations – auto3dem.pl, init_params.pm, write_params.pm, setup_rmc.pl, and findbest.pl were modified to so that calculations on a particular random model are terminated once the resolution improves to the point where the FSC curve never drops below 0.5. Also, additional random models are not calculated once one of the random models had successfully converged. In the best cases, this results in a 20x reduction in the run time relative to calculating ten full iterations for ten different random models.
- Maps generated by auto3dem have a better naming convention, using the “auto outfile” parameter to specify the file prefix and “.pif” to specify the suffix. For example, the 5th map generated where outfile equals reovirus will now be named reovirus_iter_5.pif rather than map_iter_5. Script findbest.pl modified to handle new naming conventions.
- Improved naming of variables in several of the Perl scripts and modules.
- Always output memory usage and timings in P3DR, PCTFR, PCUT, and POR. These had been under control of the hard-coded logical variables test_time and test_memory.
- Get rid of calls to bcst_parameters in Psf.f and Pcut.f. This makes the code easier to follow since bcst_parameters broadcast parameters that were listed in a common block and allows for more sensible naming of variables.
- Get rid of “include ../info.inc” in PCUT/cut_map.f and expand argument list. Get rid of “include ../mem_time.inc” in Pcut.f and declare necessary variables rather than relying on common blocks.
- Completely overhaul logic in PSF to make the sequence of operations easier to follow.

auto3dem_v1.09

- Email notification – a new module sendmail.pm was added to AUTO3DEM that provides capabilities for sending email notifications with optional attachments. This feature requires that the standard Linux/UNIX mutt email tool be installed on your system. The sendmail.pm module tests for mutt using the UNIX ‘which’ command and does a cursory check on the email address to make sure that it meets the minimal requirements for being

well formed. `auto3dem.pl` has been modified to send email notifications, with a text output of the FSC calculations, at the end of each iteration. To enable this feature, add the following line to the AUTO3DEM parameter file, where email is a valid email address.

```
auto_recipient_email
```

- Random model computations have been simplified. Modifications have been made to both `setup_rmc.pl` and `findbest.pl` so that best map is identified and copied into the directory provided as argument to `setup_rmc`. Argument list for `setup_rmc` has been expanded so that the commands file does not need to be hand edited and so that all temporary files and directories associated with random model computations have the 'RMC' prefix. This last change makes the cleanup of intermediate files more straightforward. The updated syntax for running `setup_rmc` is shown below. Note the new `ncpu` and `nodefile` arguments.

```
setup_rmc dir list nmodels ncpu nodefile [boxrad]
```

- CTF refinement – the program PCTFR (Parallel CTF Refinement) has been integrated into AUTO3DEM. This involved changes to `auto3dem.pl`, `init_program.pm`, `make_program.pm`, and `write_param.pm`. The subroutine `cmpt_ctf` called by PCTFR has been overhauled and modified so that it can handle both stigmatic and astigmatic images. PCTFR does not currently handle the transition from a stigmatic to astigmatic images unless a reasonably accurate estimate of the orientation of the astigmatism is provided. The main program `Pctfr.f` has been cleaned up and modernized.

CTF refinement is controlled by the new parameter 'auto_refine_ctf' and is enabled by default. Note that CTF refinement is only done when running in refine mode since at lower resolutions the spatial frequencies may not span enough nodes in the CTF function to make a reasonable estimate of the defocus.

- Multiplication of model projections by CTF in PO2R – In previous versions of PO2R, the projections of the model were divided by the CTF function before comparing them to the images. Modifications were made to files `Por.f` and `cmpt_ort.f` so that projections of the model are multiplied by CTF. Either approach is valid for comparing the model to the images, but multiplication by the CTF avoids the approximations that are required to properly handle the inverse CTF in the vicinity of the nodes.
- Tests added to AUTO3DEM (module `update_res.pm`) to ensure that the maximum resolution used in P3DR is not less than twice the pixel size (Nyquist limit).
- Default values now set for `PCUT_in_rad` and `out_rad`; `PPFT_annulus_low` and `annulus_high`. It should be noted though that it is still to the user's advantage to manually enter values for these parameters once estimates become available from lower resolution maps.
- PPFT modified so that existing output files are overwritten rather than having the program terminate. In addition, PPFT can now handle leading whitespace before the names of parameter files and trailing empty lines in the input file.
- General code improvements – removed unnecessary include statements in `fft_2dfft`; generalized `maptempfac` to handle arbitrarily shaped arrays; modified `readorient` to return

score and correlation coefficient data from files generated by PO2R and PPFT, respectively; modified cmpt_intrps, cmpt_ctf, and cmpt_ort to be compatible with new readorient routine;

- Temperature factor – ctf_para now applies temperature factor in the same manner (positive sign in exponent) for all values of filter. Simplifies logic of calling ctf_para with regards to sign convention. All code (except possibly PPFT) should handle temperature factor calculations correctly.
- Random model pixel binning – changed threshold for using binned data (bin_factor=2) in random model calculation from pixels of size 5Å to 4Å.
- Hollow map bug fix – Repaired bug that affects hollow map calculations when running in search mode. Since hollowed maps are an advanced feature and are normally only used in refine mode, this bug probably had no impact on any reconstructions performed to date.
- PO2R res_min bug fix – Repaired bug so that PO2R minimum resolution is properly calculated. Bug would only have been encountered if user was trying to manually override default value for res_min based on calculation involving pixel size.
- PO2R handedness test bug fix – Repaired bug that had been introduced in version 1.08 regarding test for proper hand.

auto3dem_v1.08

Modifications dealt primarily with the elimination of redundant functionality. In particular, the use of equivalent_view to generate symmetry-related orientations and crowther_to_matrix to determine rotation matrices corresponding to the orientations defined by (theta, phi, omega). These changes lead to very small numerical differences. A number of non-algorithmic modifications were also implemented that improve code readability.

- Commprk – split the file to_asym_unit.f into separate files each containing a single subroutine or function; got rid of common block and used SAVE attribute to retain the constants in function good and subroutine genrot that were previously calculated in geometry_init. Removed subroutines setrotmat and getorient since these are no longer needed. Moved the routines density_clear and maptempfac from library Compar to Commprk since they do not contain any parallel code.
- Compar – Moved routines density_clear and maptempfac to library Commprk.
- P3DR – Overhauled calculation of symmetry-related orientations for icosahedral symmetry in subroutine cmpt_intrps; replaced a fairly large block of confusing code with calls to routine equivalent_view. Replaced call to setrotmat with crowther_to_matrix.. Updated calling sequence to eight_symmetry and dihedral symmetry to take (theta, phi, omega) as arguments rather than a 3-element vector. Got rid of subroutine ico_vector since it is no longer needed for determination of equivalent orientations.
- POR – Replaced calls to setrotmat with crowther_to_matrix in subroutine cmpt_ort.
- PCTFR - Replaced calls to setrotmat with crowther_to_matrix in subroutine cmpt_ctf.

auto3dem_v1.07

- PPFT – get rid of files wavel.f, ctf_scale.f, and piraddeg.f; move ctf_firstpeak_pft.f to library Commpk, and replace calls to ctf_scale with calls to ctf_para in subroutines global_cc and calc_pfts_g. Make minor, non-algorithmic (general cleanup, redefining real arrays as complex arrays, specification of argument intents, etc.) changes to following files

| | | | |
|---------------|----------------|----------------|--------------|
| calc_pfts_g.f | fft_lop.f | global_cc.f | pftcc_res.f |
| ccf_fft.f | fft_map_fill.f | key_info.f | pftsearch.f |
| fft_2d.f | get_bestmag.f | list_ccs.f | prjavg_fft.f |
| fft_2d_back.f | get_tpo_g.f | map_fft_fill.f | |
| fft_hip.f | get_xy.f | pftcc_fill_g.f | |

- POR – change one-dimensional arrays to multi-dimensional arrays in Por.f and cmpt_ort.f. This does not change any results, but does make code easier to follow. Get rid of file imgcompctet.f and replace calls to imgcompctet with calls to imgcomport.
- P3DR - change one-dimensional arrays to multi-dimensional arrays in P3dr.F; minor modifications to matrixentries_slab.F.
- PCUT - change one-dimensional arrays to multi-dimensional arrays in Pcut.f
- PCTFR - change one-dimensional arrays to multi-dimensional arrays in cmpt_ctf.f; minor modifications to Pctfr.f
- PSF - change one-dimensional arrays to multi-dimensional arrays in Psf.f
- Compar – general cleanup of exch_3d_1.f
- Commpk – split ctf_para.f file into multiple files ctf_es.f, ctf_et.f, ctf_firstpeak.f, ctf_func.f, ctf_para.f, and ctf_temp.f. Overhaul and cleanup all routines, with an emphasis on making them usable by POR, P3DR, and PPFT. Move PPFT/ctf_firstpeak_pft.f into Commpk. Perform general cleanup of

| | | |
|----------------|---------------|--------------|
| disef_para.f | focus_astig.f | imgcomport.f |
| intlz_arrays.f | info.f | rpifimag.f |
| symmcode.f | | |

auto3dem_v1.06

- Remove files ico_EM4IMR.f and ico_EM4IMR_vector and replace with single routine in file ico_vector.f. New version may lead to very slightly different results since the hard-coded values 0.809017, 0.500000, and 0.309017 have been replaced with the more accurate representations $\cos(36^\circ)$, $\cos(60^\circ)$, and $\cos(72^\circ)$, respectively.
- Modified the code for filter=1 in ctf_para so that CTF is zeroed in when the condition $|CTF| < 0.1$ is met rather than $|CTF| < \sqrt{\text{ctf_ff2}}$. The old version could possibly lead to the CTF being set to zero for a very large range of CTF values.
- Removed the include file vax_minmax.inc since it is no longer needed by any routines.

Minor, non-algorithmic changes made to following files. Primarily removal of unused variables, specification of argument intents, improved comments, etc.

| | | |
|-------------------------|---------------------------|-------------------------|
| Commpk/arrange_3d_1.f | Ctffit/ctffit.f | P3DR/rearrange_2_slab.f |
| Commpk/cfft_1d.f | P3DR/Makefile | PCTFR/Pctfr.f |
| Commpk/ctf_para.f | P3DR/P3dr.F | PCTFR/cmpt_ctf.f |
| Commpk/fft_2dfft.f | P3DR/cmpt_intrps.F | POR/Por.f |
| Commpk/imgcomport.f | P3DR/eight_symmetry.f | POR/cmpt_ort.f |
| Commpk/interpl_3d.f | P3DR/exch_intp.F | POR/imgcompct.f |
| Commpk/intlz_params.f | P3DR/exch_intpol.F | PPFT/ccf_fft.f |
| Commpk/pfile_info.f | P3DR/fftsynth_1_m_slab.f | PPFT/pftcc_res.f |
| Commpk/readorient.f | P3DR/fftsynth_2_m_slab.f | PSF/Psf.f |
| Commpk/to_asym_unit.f | P3DR/ico_vector.f | PSF/comp_sfactor.f |
| Compar/density_clear.f | P3DR/matrixentries_slab.F | include/info.inc |
| Compar/exch_3d_1.f | P3DR/move_data_2_slab.f | include/mem_time.inc |
| Compar/output_density.f | P3DR/realtocomplx.F | |

auto3dem v1.05

- Improved capabilities added for exiting auto3dem if one of the image processing programs aborts. Depending on operating system and batch queuing system, the return value from launching an MPI job with system command may not be accessible to the auto3dem script. To ensure that the abort is detected, the output from the MPI program is parsed and auto3dem exits if the string MPI_Abort (not case sensitive) is found. Involved modifications to error_stop.f, run_mpi_prog.pm, and auto3dem.pl.
- Minor performance improvements made to 3D interpolation routine interpl_3d.
- General cleanup of cmpt_intrps.F, fftsynth_1_m_slab.f, exchange_2_slab.f, P3dr.F, Pcut.f, ctf_para.f, Psf.f, Por.f, Pctfr.f, com_01.inc, intlz_arrays.f, and intlz_params.f. Specification of argument intents, removal of unused variables, etc.
- General cleanup of PSF/comp_sfactor.f. Move calculations contained within subroutine cplex2ap into comp_sfactor and get rid of unnecessary operations.
- New control parameters added:
 - auto switch_mode (see note below)
 - auto term_refine (functionality not yet active)
 - auto term_search (functionality not yet active)
- Auto3dem now monitors resolution and can automatically switch from search to refine mode if the following conditions are met:
 - switch_mode flag is true
 - Resolution has not improved by at least 0.25Å over previous iteration
 - Mode is currently search
- Iteration information is now tracked so that numbering from one run to the next is maintained. Restart and continuation files specify the last iteration that had been completed. The restart files also keep track of the number of iterations required to finish the original calculation.
- For restart or continuation, new results are appended to the log and summary files.
- Restart files are written at two points in each iteration: after the completion of origin and orientation refinement and after construction of map. The restart files properly set the auto_have_map flag so that map is constructed from particle images if program needs to be restarted after first checkpoint.
- Output files are named consistently using the name of the auto outfile parameter:
 - outfile_log – detailed output log
 - outfile_summary – summary information
 - outfile_restart_na – restart after first checkpoint, iteration *n*

- o outfile_restart_nb – restart after second checkpoint, iteration n
 - o outfile_continue – continuation file for successfully completed runs
- Summary file now lists the number of particles that had been selected to construct the final map and the number of CPUs (MPI processes) used.
- Format of restart and continuation files has been improved to logically group parameters.
- Default value (0.1 pixels) specified for po2r dcenter, step size for origin refinement.
- auto final_map no longer used and has been labeled as deprecated.

auto3dem_v1.04

- Auto3dem version now printed in log and summary files
- Symbolic links to executable Perl scripts in BIN directory created. If path information is setup correctly, can use name of script with or without .pl extension.
 - o auto3dem → auto3dem.pl
 - o setup_rmc → setup_rmc.pl
 - o handflip → handflip.pl
 - o findbest → findbest.pl
- auto3dem and setup_rmc executed without any arguments provides usage information and version
- Improved format for auto3dem summary data
- Fixed bug in auto3dem.pl where PPFT output files were moved before they were no longer needed.
- Minor changes to output from update_res.pm
- Default value for lower resolution limit of FSC curve (psf res_min) set to 60Å
- Default value for PO2R angular step size (po2r dangle) set to one-half of the angular step size for PPFT.
- Default value for PO2R res_min set to $2/5 * \text{boxrad} * \text{pixelsize}$
- Default value for PPFT resolution_low set to $2/5 * \text{boxrad} * \text{pixelsize}$
- Test provided on number of iterations in input file. Must be defined and be at least zero.
- Test added to make sure that at least one 'data' line is specified and that if wildcards are used they expand to include at least one file.

auto3dem_v1.03

Changes were implemented that make it much easier to run AUTO3DEM. Rather than creating symbolic links to Perl code and binaries, path information is simply added to the computing environment. This required some changes to the Perl code so that binaries and custom modules would always be found.

- **AUTO3DEM:**
 - o Custom Perl modules reside in the same directory as the executable Perl scripts. Added the following line to all .pl and .pm files so that custom modules would always be found, regardless of where the scripts are executed


```
use lib do { __FILE__ =~ m|^(\.?.?)[^/]*$|; "$1"; };
```
 - o Modified setup_rmc.pl so that executable Perl scripts in the command file are not pre-appended with “./”. For example

./auto3dem.pl → auto3dem.pl

- In module init_params.pm, get rid of full path information for binaries
- **make_all:** Modified so that all binaries are copied into the BIN directory

auto3dem_v1.02

- **AUTO3DEM:**
 - Modified auto3dem.pl and init_params.pm so that map can be constructed without doing resolution estimation. This feature is particularly useful when applying inverse temperature factor to final map.
 - Particle selection criteria can now be applied globally across all files or on a per file basis. By default, selection criteria applied globally. Choice controlled by auto_global_select.
 - Create subdirectory for storing the program input and output files, together with the filtered particle parameter files generated by applying the particle selection criteria.
 - Rename the PSF output file corresponding to the best particle selection criterion for a given iteration to FSC_curve_{iteration}.
 - Handle the case where the FSC curve never drops below fsc_locut. Particular useful in the early stages of search mode where resolution of map has the most rapid improvements.
 - Get rid of dryrun option
 - Write summary file containing just the most important information required to monitor the progress of the reconstruction.
 - Modify init_params.pm to ignore 'auto dryrun' and 'auto split_ppft' and provide message that these are deprecated options.
- **Program POR:** Modified Por.f so that program can deal with map files that contain ridiculous values for the pixel size. In this case, the pixel size from the particle parameter file is used instead.
- **handflip.pl:** Script for changing the handedness of orientations in particle parameter files.

auto3dem_v1.01

- **Program P3DR:** Interpolation routine matrixentries_slab.F improved to yield better performance. Actual speedups will vary by problem and hardware, but runs using test problem show 28% reduction in time spent in 2d interpolations and a 15% reduction in overall run time.
- **Program P3DR:** Modified P3dr.F and get_1st_ortid.f to properly handle empty particle parameter files.
- **Program PPFT:** Added option for verbose=-1, which results in the correlation coefficient cc_cmp not being calculated. This option should only be used for construction

of starting model using random model method since cc_cmp is generally used to select particle images.

- **AUTO3DEM random model method:** Script setup_rmc.pl modified so that default values for PPFT are overridden.
 delta_theta = 1° (default 0.5°)
 verbose = -1 (default 2)
- **AUTO3DEM:** parameter file now accepts po2r for specifying PO2R input. por still works to allow back compatibility with v1.0, but all internal perl variables renamed to use po2r (e.g. %por_ref → %po2r_ref)